

## Medex Advanced Multimedia File Analysis Training

### DAY 1

#### Introduction

- Medex Instructors, local attendees
- Scope and objectives of training course
- Review of provided documents and samples

#### 1.1 Bitwise and Byte-wise information representation

- Introduction to digital files in general
- Specifically
  - Bits
  - Bytes
    - Understanding Binary
    - Understanding Hexadecimal
  - Files
    - Stored segment or block of information available to a computer program
  - Interpreting Files
    - File specifications
    - ASCII
- Generally
  - Header
    - Metadata
  - Payload
    - Content
- File systems and Metadata

#### 1.2 Overall Multi-media format structural design (using ISO Base Media files - MP4, MOV, 3GP, 3G2, M4V - as example)

Goal: Understanding the dependencies of all top-level components of ISO Base Media files.

- MP4 Concepts
  - Presentation
  - Streams
  - MP4 Temporal structure
    - Timescale
  - MP4 Physical structure
    - Boxes
  - Format Specs/video decoding
    - ISO BMFF
    - Demonstration of decoding
  - Manual decoding assignment
  - Assumptions for this course

- Ignoring audio codecs for now

### 1.3 Critical ISO Base Media boxes

- ftyp, moov, mvhd, trak, tkhd, mdia, minf, vmhd, dinf
- Stbl
  - stsd--->avcC for parameters
  - Main important Sample tables
    - stsz - sample sizes
    - stsc - samples to chunks
    - stco - chunk offsets
    - stts - sample decode times
    - cts - sample composition times

## LUNCH

### 1.4 How samples are stored in MP4

- mdat
  - Interleaved
  - In chunks

### 1.5 How h264 is structured (generally) - NALUs

- NALUs
  - NALU Structure
  - NALU Types
- Storage Implementations
  - Annex B
    - Start Codes
    - Emulation prevention bits
  - AVCC
    - “Packed” in MP4
  - AVC1 (h264) example
    - Samples as Access Units
    - Stored in mdat
    - Stored without start codes
    - Stored with size length bytes at beginning of each Access Unit
    - stds parameters info - pps and sps stored separately in moov

## DAY 2

### Introduction

- Review of Day 1 topics and activities, outstanding questions
- Scope of Day 2

### 2.1 Exercises: Manual Approaches to File Repair

Goal: Review of possible scenarios that require file repair on ISO Base Media Format files. Understand when and how a file can be repaired, and when it cannot.

- How can files be broken?
- Manual approaches to repair - when MP4 file is present but broken
  - Missing metadata
  - Incorrect metadata
  - Structural inconsistencies

## LUNCH

### 2.2 Exercises: Automated Approaches to File Repair

- Approaches to repair - when only mdat data is present - Decision tree
  - Considerations for reconstituting mdat video data
- Short Introduction to Python
  - IDE
  - Basics
  - Data Types
  - Reading and Writing binary files
  - Resources
- Using code and logic to script automated restoration
  - AVCC to Annex B example
  - Walkthrough code
- 5 Test Examples
  - Complete mdat with no audio, use sps/pps from file from same source
  - Complete mdat with no audio, use sps/pps from file from different source
  - Partial mdat with no audio, use sps/pps from file from same source
  - Complete mdat with audio, use sps/pps from file from same source
  - Random run of bytes, mdat with or without audio, use sps/pps from file from different source

## DAY 3

### Introduction

- Review of Day 2 topics and activities, outstanding questions
- Scope of Day 3

### 3.1 - Continued learning on File Format Structure (AVI, WAV)

- File format specifications (RIFF/AVI, RIFF/WAV)
  - File structure
  - Format decoding
  - Manual decoding assignment

### 3.2 - Validating Metadata Extraction tools

- What is metadata extraction?
  - Binary data
  - File specification
  - Human readable reporting
- Automated methods
- File system options
- Review and compare output of common extraction tools
  - MedialInfo
  - Exiftool
  - FFprobe
- Exercises
  - Based on learning from Day 1, we will manually validate the output of automated metadata extraction tools.

## LUNCH

### 3.3 What is Video Authentication

- Walkthrough of authentication case using current approaches
- Group discussion of how we would approach authenticating a couple of example videos
- How we would do it in Medex

### 3.4 - Structural Analysis and Classification Methods

- Structural analysis
  - Structural comparison group activity
- Supported file types
- Reference library
- Classification
- Overview of the Medex UI and data results
  - Submitting files
    - Direct Web App
    - Local Processing Client
    - Hands-on file submission
  - Working with projects
  - Project navigation
  - Working with files - similar to intro to show layout
  - General info
  - Structural analysis
  - Structural signature
  - Classification
  - Metadata analysis
  - Proprietary Metadata
  - Interactive structural view
  - Manipulation/Validation
  - Notes

- Reporting

## DAY 4

### Introduction

- Review of Day 3 topics and activities, outstanding questions
- Scope of Day 4

### 4.1 - Interpreting Results

- Identification of the question being addressed
- Structural analysis
  - Brand
  - Make/Model
  - Generational history
  - Signature matches
  - Non-matches
  - What we can say about a file when excluding others in ref library (like editing)
  - Exercises
- Classifiers results
  - Brand
  - Make/Model
  - Exercises

### 4.2 Interpreting Results - continued

- One to one comparison
  - Using Exemplars
  - Exercises
- Common encoding themes (with examples)
  - FFmpeg
  - Apple
  - Samsung
  - Adobe
  - Exercises
- Using proprietary structure outputs
  - Exercises
- Using interactive viewer
  - Exercises
- Using semantic metadata in context/contrast with Medex results above
  - Exercises
  - Apple trimmed on device example

## LUNCH

### 4.3 - File Modification/Verification Tests

- Difference between alteration and manipulation of hex data
- Interpreting failed tests
- Known encoding issues (i.e. WhatsApp)

#### 4.4 - Using structural analysis for deepfake analysis

- Strengths of this approach
- Weaknesses of this approach

#### 4.5 Hands-on Practical

- Expressing opinions/results
- Practical assignments

#### 4.6 Certification of Completion for all attendees

[Optional]

#### 4.7 Sit for Medex Certification Test (1 hour)

- Online google form
- 20 questions
- Certified Medex Examiner